



JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

MACHINE LEARNING BASED ALGORITHMIC APPROACH for BITCOIN PRICE PREDICTION

- #1 G Balamurugan , Asst Prof , Mohamed Sathak Engineering College , Kilakarai /
 #2 Mr. D.Sheik Mohamed Rashid , Mohamed Sathak Engineering College , Kilakarai /
 #3 Dr.S Sajitha Banu, Asst Prof , Mohamed Sathak Engineering College , Kilakarai
 #4 M Mohamed Rafi , Professor & Head , Mohamed Sathak Engineering College , Kilakarai /
 #5 K.Ramya , Asst Prof , Mohamed Sathak Engineering College , Kilakarai

Machine learning methods have been applied in many fields, including finance. Risk and sentiment analysis in particular can benefit from algorithmic machine learning models and their ability to predict the future price of securities. This article presents our experience in applying machine learning methods to analyze cryptos like bitcoin or ether, which represent an entirely new asset class with no historical data at all. We discuss our methodology for evaluating risk for these instruments, which may be useful for investors who are new to cryptocurrencies. We also take look at the differences between traditional equity- and bond-based markets, as well as emerging crypto markets, as they relate to financial risk measures and how we can apply them to predict future price movements.

Keywords – Bitcoin, Machine Learning, Prediction, ML Algorithms, Python

I. INTRODUCTION

Cryptocurrencies have proven to be a part of the global economy as they grow and develop each day, providing different secure features to the users and users. The future of cryptocurrencies looks secure and promising, looking at how much more they are limited in production, being more secure than most other forms of currency and they contain less risk. The cryptocurrency market is a relatively young sector of the financial industry, with Bitcoin being the oldest, largest and most valuable of all digital currencies. But there are now more than 800 cryptocurrencies, many of which present unique capabilities and use cases. In this age of digital transformation, investors would be wise to review and consider these six blockchain-based cryptocurrencies before making an investment in 2019 and beyond.

Bitcoin makes up 62% of all cryptocurrencies
 Total capitalization of traded virtual currencies

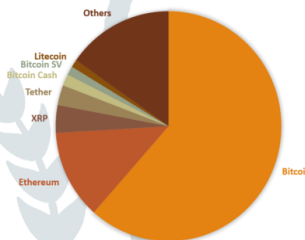


Figure 1. Cryptocurrency Market Share worldwide

Bitcoin is a revolutionary and decentralized worldwide digital currency. It can be used by anyone and is marketed over the internet. The community that regulates the production of Bitcoin consists of individuals and organizations around the world who arrange exchanges using programming code to approve different upgrades and so control the production of new Bitcoins. As of now, Bitcoin has acquired financial backers and organizations as it has encouraged services and product deals. In addition, Bitcoin has also made itself the foremost authority in decentralized digital currency. It is completely digital, and there are no actual paper bills. The system is based on the fact that you own private keys, which can be used to exchange currency or credit other accounts. You can send or spend your bitcoins from anywhere in the world without incurring any transaction fees.



Figure 2. Regionwise market of bitcoin (Volumes)

There has been extensive research and paper work done on bitcoin. Predicational papers on the same is of very limited scope. This is due to the complex computation process and the extensive work that has to be done to retrieve the datasets that has been accumulated over the years. The market of the Bitcoin has been roaring high as per the world-wide market price of bitcoin has soared from \$124 to \$34k in a span of 10 years. (Ref Figure 3. Chart display of Bitcoin market price)

This paper is intended to give a lucid picture of the bitcoin pricing. We aim to develop an efficient model aimed towards predicting the future value incentive of bitcoins in addition to that we endeavor to disclose the shortcomings being faced by statistical analysis techniques and also lay down some recommendations for development in this field.



Figure 3. Chart display of Bitcoin market price

The price of an asset is a reflection of the value that the market participants attribute to it. Bitcoin is an asset whose price has witnessed wide fluctuations in interest, visibility and engagement of investors and traders. The popularity of bitcoin has seen a phenomenal rise over the past few years, giving it a unique place in the global financial ecosystem.

Our model uses recent bitcoin trading data to predict future bitcoin prices. I collect data from the largest bitcoin exchange and, using descriptive statistics and different machine learning models, try to find which indicators are the best predictors of bitcoin prices. Hence a rapid , trustworthy and highly efficient system with using the latest machine

Learning algorithms and with a ease of use GUI is the need of the hour . That system must be capable of taking into the consideration of the years of data and market fluctuations to account , so that the model will be of high accuracy.

II. EXISTING SYSTEM

Intelligent pricing prediction models can be developed by integrating artificial intelligence (AI) and machine learning (ML) techniques with the recent advances in high performance computing. In a scoping study, we review the current literature on predictive ML/AI systems that support pricing decisions.

The comparison between time series and cross-sectional data is not enough to perform a similar forecasting model analysis. This is because most of the time series are typically heteroskedastic, while many of the cross-sectional data might be heteroskedastic or homoskedastic--this depends on whether the sample is a population or a random sample. Thus, an accurate prediction model of bitcoins based on econometric methods can only be made if the stochastic nature of both time series and cross-sectional data is considered.

There has been some previous works and calculated research done on the prediction of digital currencies and bit coin. Earlier works use the algorithms that take only the older datasets. Recent and current running datasets cannot be fed into the prediction system. Moreover, the prediction is based on MATLAB and hance the computational time is higher. Principal computation time must be rather very less so that the precision calculation time is lesser and more accurate.

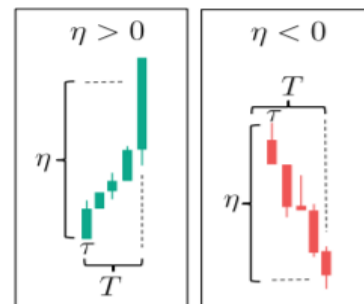


Figure 4. Existing Models Bitcoin price rise or fall over N

Predicting the cost of Bitcoin with a machine learning methodology is complex. Existing systems rely on comprehensive analysis, but over a period of time and only in terms of market fluctuation, specifications, and accuracy.

Some of our research has been implemented in MATLAB and if you'd like to use it to predict cryptocurrency prices with accuracy over a certain period of time. There are some accuracy levels, but they fall short of what financial professionals would say is accurate enough. Finance professionals would like to have more precise predictions than what is currently available. Some research has been done on cryptocurrencies. But, little work has been done to predict the price of the cryptocurrency over a certain period of time.

III. PROPOSED SYSTEM

Our goal is to create a model that would make accurate predictions on a particular pricing for the bitcoin currency. For this purpose, we will use data from the public domain bitcoin forums. We'll rely on self-reported information given by users, i.e. their experience and opinions about the bitcoin price at which they sold any of their previously owned bitcoins. Machine learning is a field of technology which deals with development of computer programs that can learn from data without being explicitly programmed. The programs are capable of getting better results as the amount of data used for training the program increases.

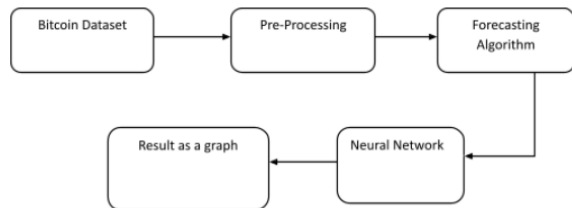


Figure 5. Process flow Diagram of the proposed method

This paper introduces an innovative way to predict the price of bitcoins. It uses multiple machine learning algorithms including the Random Forest and Support Vector Machine algorithm. The system is very accurate, and it can predict bitcoin prices with an accuracy of %96 for up to 5 days.

We have to build a system that takes large amounts of historical Bitcoin data, process it with Machine Learning algorithms, prunes the data we don't need, then predicts what the Bitcoin price might be in the future. We will keep track of our predictions, and our algorithm will get better as it is used more often.

A. DATA COLLECTION:

We are taking inputs of real numbers and integers. The procedure is implemented in python idle so that accuracy levels can be increased.

The steps that are done are Preprocessing, Desired Feature Extraction for both the input image and dataset images so as to increase the accuracy level of prediction.

B. PREPROCESSING & EXTRACTION

With the large data generated at the almost all sources in recent times, it is mandatory to process and analyze them. The preprocessing of the data is one of the initial steps towards the predictive data processing concepts. The datasets are preprocessed using the DWT Algorithm by applying filtering techniques. The decision tree algorithm helps in extracting some new features which are not present in the original data set that can be accumulated over time.

Feature extraction is a way to use the information in datasets to decrease the number of highlights, which allows for more variety. It will work to decrease the number of highlights in a dataset by making new features from the current ones. This is done by using the decision tree algorithm.

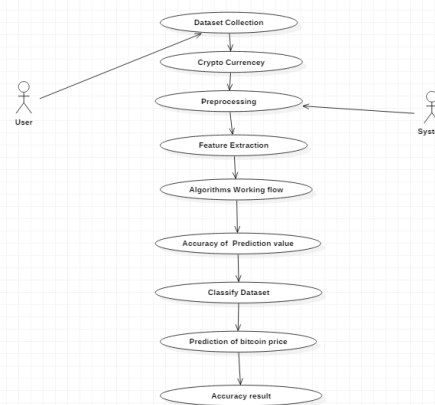


Figure7. Usecase Diagram of the Bitcoin Prediction

C. PREDICTION

There are many online data services that provide you the option to purchase the data to download. Such options offer a vast range of datasets, suitable for a variety of purposes, however not all of these datasets adhere to the strictest of standards and can be sometimes unreliable. Dataflow analyses our data with respect to specific technical norms such as reliability, usability, relevance and what it is capable of achieving such as prediction. We attempt to identify the needs of our project in order to evaluate if the dataset is ideal for your requirements and then establish the true value and pricing of such a prediction model

D. ACCURACY ANALYSIS

The results obtained has to be made comparable with the existing data, so that we can clearly manipulate the context and the way of the prediction algorithm to gain a more indepth categorization of the prediction models. The Results of the same are stored into the database for further detection and analysis purpose. This will help to increase the accuracy of the prediction model and will gain self sufficient mode of the prediction model. The algorithmic approach with the best level of accuracy can be trained with multiple sets of data and model can be established as a convenient working model.

market	rpt_key	last	diff_24h	diff_per_2bid	ask	low	high	volume	
bitstamp	btc_eur	1996.72	2029.99	-1.63892	2005.5	2005.56	1950	2063.73	2314.501
bitflyer	btc_jpy	267098	269649	-0.94604	267124	267267	267124	267267	70922.88
korbit	btc_krw	3003500	3140000	-4.34713	3003500	3004000	3002000	3209500	6109.753
bitstamp	btc_usd	2237.4	2239.37	-0.08797	2233.09	2237.4	2154.28	2293.46	13681.28
okcoin	btc_usd	2318.82	2228.7	4.043613	2319.4	2319.99	2129.78	2318.82	4241.642
korbit	etc_krw	22740	23150	-1.77106	22700	22730	21000	25500	855853.4
bitflyer	eth_btc	0.1034	0.08855	16.77019	0.10315	0.1034	0.10315	0.1034	21670.57
korbit	eth_krw	311800	274500	13.58834	311800	311950	272500	336000	327416.9
bitflyer	fx_btc_jpy	266600	275066	-3.07781	266276	266640	266276	266640	70921.73
okcoin	ltc_usd	25.42	24.931	1.961414	25.364	25.42	23.2	25.514	171028.4
korbit	etc_krw	22740	23010	-1.1734	22700	22730	21000	25500	855853.4
korbit	eth_krw	311900	272500	14.45872	311900	311950	273000	336000	326531.2
bitstamp	btc_eur	2005.56	2013.41	-0.38989	2005.56	2006.01	1950	2063.73	2317.22
bitflyer	btc_jpy	268271	269440	-0.43386	268271	268300	268271	268300	71179.52
korbit	btc_krw	3003500	3140000	-4.34713	3003500	3004000	3002000	3209500	6116.842
bitstamp	btc_usd	2248.39	2242.44	0.265336	2247.77	2248.38	2154.28	2293.46	13701.7
okcoin	btc_usd	2320.42	2228.4	4.12942	2320.99	2321.49	2129.78	2322	4260.262
bitflyer	eth_btc	0.1034	0.08813	17.32668	0.10315	0.1034	0.10315	0.1034	21524.42

Figure 6. Input samples Dataset



Figure 8. Realtime Vs Predictive Price of Bitcoin

IV. SOFTWARE SYSTEM

A. Python 2.5 / 3.5

This paper is implemented using Python. Python is deployed in the paper because of the extensive libraries it has for the calculations purpose and its ability to handle large amount of data with a minimal processing speed. Extensive libraries available for the data processing tasks makes it easy for the programmers and alternative users to understand the extensivity of the code and the process flow. Prediction and analysis of the same can be done with ease using the python flow.

B. Anaconda Navigator

Anaconda Navigator is the graphical user interface to load the python program and view the visualizations. We will be able to manage the packages and running environments using the Anaconda Navigator with ease.

V. RESULTS & CONCLUSION

Thus, a statistical model for the prediction of Bitcoin prices at random is done with a higher precision of accuracy. With a substantial increase in the datasets and with lesser processing time, this model will be self-sufficient and with optimum operational costs. Efficacy of the algorithm is on the high so that the preconditional value of the prediction can be computed with ease.

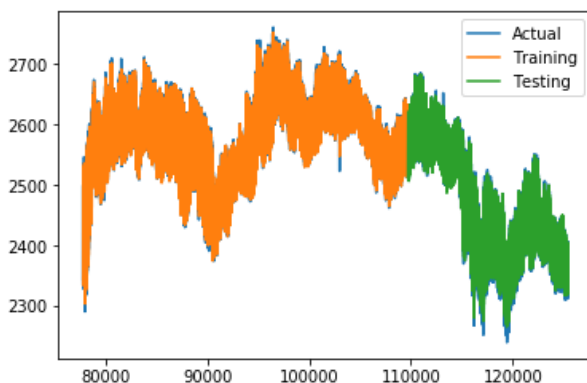


Figure 9. Bitcoin Pricing prediction results

VI. FUTURE SCOPE OF WORK

This paper implements the prediction model of the bit coin with a highly self-sufficient and enhanced accuracy level. The scope of the work can still be increased if there are large volumes of data that can be fed in to the system with the number of years increased. This will create a perfect model for the prediction analysis. Latest algorithms and application tools can also be deployed so that the work of the model is proven to be good.

REFERENCES

- [1]. S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008.
- [2] G. Heleman, M. Rauchs, "Global cryptocurrency benchmarking study," Cambridge Centre for Alternative Finance (2017).
- [3] G. Neil, H. Halaburda. "Can we predict the winner in a market with network effects? Competition in cryptocurrency market," Games, vol.7 no.3 , 2016, pp. 16.
- [4] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. Vanderplas, A. Joly, B. Holt, and G. Varoquaux, "API design for machine learning software: experiences from the scikit-learn project," arXiv preprint arXiv:1309.0238, 2013.
- [5] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G S. Corrado, A. Davis, J. Dean, M. Devin, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu and X. Zheng "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," arXiv preprint arXiv:1603.04467, 2016.
- [6] C. François, "Keras: Deep learning library for theano and tensorflow," URL: <https://keras.io/k>, 2015.
- [7] D. Shah, K. Zhang, "Bayesian regression and Bitcoin," Communication, Control, and Computing (Allerton), 2014 52nd Annual Allerton Conference on. IEEE, 2014.
- [8] I. Madan, S. Saluja and A. Zhao. "Automated Bitcoin Trading via Machine Learning Algorithms," 2015.
- [9] A. Greaves, B. Au, "Using the Bitcoin Transaction Graph to Predict the Price of Bitcoin," 2015.
- [10] J. Almeida, S. Tata, A. Moser and V. Smit, "Bitcoin prediction using ANN," June 2015
- [11] S. McNally, "Predicting the price of Bitcoin using Machine Learning" Diss. Dublin, National College of Ireland, 2016
- [12] S. Patro, K. Sahu. "Normalization: A preprocessing stage," arXiv preprint arXiv:1503.06462, 2015.

[13] X. Dang, H. Peng, X. Wang and H. Zhang "Theil-Sen Estimators in a Multiple Linear Regression Model," Olemiss Edu, 2008.

[14] P. Huber. "Robust statistics." International Encyclopedia of Statistical Science. Springer Berlin Heidelberg, 2011, pp. 1248-1251.

[15] A. Bernal, S. Fok and R. Pidaparathi. "Financial Market Time Series Prediction with Recurrent Neural Networks." 2012.

[16] S. Hochreiter, J. Schmidhuber. "Long short-term memory." Neural computation vol.9 no.8, 1997, pp. 1735-1780.

[17] K. Cho, B. Van Merriënboer, C. Gulcehler, D. Bahdanau, F. Bougares, H. Schwenk and Y. Bengio "Learning phrase representations using RNN encoder-decoder for statistical machine translation." arXiv preprint arXiv:1406.1078, June 2014

[18] Greaves A, Benjamin A (2015), Using the bitcoin transaction graph to predict the price of bitcoin.

[19] Madan SS, Zhao A (2015), Automated bitcoin trading via machine learning algorithms.

[20] Hitam NA, Ismail AR (2018), Comparative Performance of Machine Learning Algorithms for Cryptocurrency Forecasting. Indonesian Journal of Electrical Engineering and Computer Science, 11(3).

[21] Guo T, Antulov-Fantulin N (2018), Predicting short-term bitcoin price fluctuations from buy and sell orders. arXiv preprint arXiv:1802.04065.

[22] Jang, H., & Lee, J. (2018). An empirical study on modeling and prediction of bitcoin prices with bayesian neural networks based on blockchain information. IEEE Access, 6, 5427-5437

